ARMY RESEARCH LABORATORY

Prediction of Protein-Peptide Interactions: Application of the XPairIt API to Anthrax Lethal Factor and Substrates

by Margaret M. Hurley and Michael S. Sellers

ARL-RP-456 **September 2013**

A reprint from *Proceedings of SPIE*,
Vol. 8719, 87190A, 2013.

Approved for public release; distribution is unlimited.

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

Army Research Laboratory
Aberdeen Proving Ground, MD 21005-5069

ARL-RP-456 **September 2013**

Prediction of Protein-Peptide Interactions: Application of the XPairIt API to Anthrax Lethal Factor and Substrates

Margaret M. Hurley and Michael S. Sellers
Weapons and Materials Research Directorate, ARL

A reprint from *Proceedings of SPIE*,
Vol. 8719, 87190A, 2013.

Approved for public release; distribution is unlimited.

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED (From - To)
September 2013	Reprint	1 October 2012–17 April 2013

4. TITLE AND SUBTITLE
Prediction of Protein-Peptide Interactions: Application of the XPairIt API to Anthrax Lethal Factor and Substrates

5a. CONTRACT NUMBER
ORAUW911QX-04-C-0129

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Margaret M. Hurley and Michael S. Sellers

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Research Laboratory
ATTN: RDRL-WML-B
Aberdeen Proving Ground, MD 21005-5069

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-RP-456

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
A reprint from *Proceedings of SPIE*, Vol. 8719, 87190A, 2013.

14. ABSTRACT
As software and methodology develop, key aspects of molecular interactions such as detailed energetics and flexibility are continuously better represented in docking simulations. In the latest iteration of the XPairIt API and Docking Protocol, we perform a blind dock of a peptide into the cleavage site of the Anthrax lethal factor (LF) metalloprotein. Molecular structures are prepared from RCSB:1JKY and we demonstrate a reasonably accurate docked peptide through analysis of protein motion and, using NCI Plot, visualize and characterize the forces leading to binding. We compare our docked structure to the 1JKY crystal structure and the more recent 1PWV structure, and discuss both captured and overlooked interactions. Our results offer a more detailed look at secondary contact and show that both van der Waals and electrostatic interactions from peptide residues further from the enzyme's catalytic site are significant.

15. SUBJECT TERMS
anthrax lethal factor, peptide binder, flexible docking, in silico design

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Margaret M. Hurley
Unclassified	Unclassified	Unclassified	UU	16	19b. TELEPHONE NUMBER *Include area code* 410-306-0728

Standard Form 298 (Rev. 8/98)
Prescribed by A SI Std. 39.18

Prediction of Protein-Peptide Interactions: Application of the XPairIt API to anthrax lethal factor and substrates

Margaret M. Hurley* and Michael S. Sellers
U.S. Army Research Laboratory, Weapons and Materials Research Directorate, Aberdeen Proving Ground, MD, USA 21005

ABSTRACT

As software and methodology develop, key aspects of molecular interactions such as detailed energetics and flexibility are continuously better represented in docking simulations. In the latest iteration of the XPairIt API and Docking Protocol, we perform a blind dock of a peptide into the cleavage site of the Anthrax lethal factor (LF) metalloprotein. Molecular structures are prepared from RCSB:1JKY and we demonstrate a reasonably accurate docked peptide through analysis of protein motion and, using NCI Plot, visualize and characterize the forces leading to binding. We compare our docked structure to the 1JKY crystal structure and the more recent 1PWV structure, and discuss both captured and overlooked interactions. Our results offer a more detailed look at secondary contact and show that both van der Waals and electrostatic interactions from peptide residues further from the enzyme's catalytic site are significant.

Keywords: anthrax lethal factor, peptide binder, flexible docking, *in silico* design

1. INTRODUCTION

The prediction of interactions between a protein and a peptide substrate present a unique challenge to the computational community[1]. Incorporation of flexibility (both target and substrate), solvent effects, and detailed energetics play an important role in the description of this contact, as well as a means to enhance it. We have recently reported progress in the development of peptide binders for the anthrax (Bacillus anthracis) protective antigen in a combined experimental and computational effort to develop improved bioreceptor elements for sensing applications[2, 3]. As part of our continued computational effort to analyze, understand, and *a priori* predict the assembly of protein-peptide substrate complexes, we turn our attention to a related system, the anthrax lethal factor.

Anthrax lethal factor (LF) is one of three individual proteins that together compose anthrax toxin. LF is a Zn2+ metalloprotein which has been shown to interfere with cellular signaling pathways through cleavage of mitogen-activated protein kinase kinase (MAPKK) enzymes at their N terminus. Alone, LF is not toxic. A crystal structure of the anthrax lethal factor has been available for over a decade[4], greatly facilitating the experimental and computational study of structural factors leading to selectivity in peptide inhibitors of the LF protein. LF is composed of 4 domains (shown in Figure 1a) with binding and cleavage of MAPKK proteins occurring in a long groove along the interface between Domains II, III, and IV. The zinc-containing active site contiguous to this groove is located within domain IV. Mutagenesis studies demonstrate that Domain III is essential for activity and is believed to play a role in fixing the substrate in place prior to scission[4]. Turk et al[5] applied a mixture-based peptide library and solved multiple LF-substrate crystal structures to determine characteristics of an optimal peptide binder. Importantly, they found a bound conformation somewhat closer to the active site and with orientational differences relative to the 1JKY structure of Pannifer et al., using a similar (but not identical) peptide. Due to the high cleavage efficiency of LF for a number of these peptides, crystal structures were also obtained in the absence of the catalytic Zn, allowing accurate determination of the structure of the pre-cleavage complex. In their analysis, Turk et al. postulate the hydrophobic nature of the groove to be a primary factor determining selectivity. Groove hydrophobicity is visible in Figure 1b, where the LF residues are colored by the Eisenberg scale. We note, however, the groove has some mixing of hydrophobic character, and may not wholly be a sufficient driving force for binding. Hong et al.[6] used a variety of computational techniques to study the LF-peptide substrate interaction. They found that van der Waals interactions are important to stabilize the substrate, as well as noting the importance of substrate-LF hydrogen bonds in Tyrosine-containing peptides. Dalkas et al[7] used docking and molecular dynamics simulations to study a variety of MAPKK-based peptide inhibitors and their interaction

*Margaret.m.hurley12.civ@mail.mil; phone 1 410 306-0728; fax 1 410 306-1909

Smart Biomedical and Physiological Sensor Technology X, edited by Brian M. Cullum, Eric S. McLamore,
Proc. of SPIE Vol. 8719, 87190A · © 2013 SPIE · CCC code: 1605-7422/13/$18 · doi: 10.1117/12.2014767

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/31/2013 Terms of Use: http://spiedl.org/terms

with LF. They were able to tabulate LF residues exhibiting major electrostatic interactions with these inhibitors. This work also stressed the importance of van der Waals interactions in substrate binding, and determined that LF Tyr728 is essential to activity for stabilization of the scissile bond, possibly through the backbone carbonyl group.

In order to analyze the details of peptide binding to the anthrax LF, we apply our XPairIt API[3, 8] to the case of the LF protein interacting with a 16-residue peptide emulating the MAPPK-2 N-terminus. To probe the more subtle aspects of this interaction, particularly interactions within the groove which may contribute to formation and stabilization of a precleavage complex, this is performed in the absence of the catalytic Zn. Results are then examined in the context of the 1JKY crystal structure, as well as the 1PWV structure of Turk for the similar LF-LF20 peptide complex, both of which are Zn-free. The XPairIt API, in addition to combining powerful docking (PyRosetta)[9] and molecular dynamics (NAMD) software[10], enables a very detailed level of analysis on the resulting structure to sift through and highlight, on a per-residue basis, the forces driving formation of the protein-peptide complex. The efficacy of a multi-method approach to flexible docking has been demonstrated in numerous places[11-17]. We note that previous computational studies mentioned above have highlighted the importance of van der Waals interactions in these particular systems. With this in mind, we augment our analysis by providing docking scoring results for vital interactions within the context of results from recently developed software suited to characterizing non-covalent interactions, the NCIplot program of Johnson et al.[18, 19] In addition to providing a visual representation of the complicated web of interactions in a protein-ligand complex, this software also enables the characterization and quantification of these interactions, thus allowing an additional tier of validation on the energy criteria used to generate these complexes.

2. METHODOLOGY

In this section, we explain our methods of structure preparation and equilibration, using several external programs connected through the XPairIt API. Additionally, we summarize the specialized *XPairIt Docking Protocol* used in the global docking search of the MAPKK-2 peptide and Anthrax Lethal Factor (LF) protein, and the techniques used in analysis of the 1JKY crystal structure and our docking results.

2.1 Preparation and Equilibration of the LF Protein and MAPKK-2 Peptide

As obtained from the RCSB Protein Data Bank, the 1JKY structure comprises the Anthrax Lethal Factor (LF) protease with a Zn-free active site, in complex with the N-terminus 16-residue peptide of a mitogen-activated protein kinase kinase 2 (MAPKK-2) at a resolution of 3.90Å. The crystal structure of the LF protein in 1JKY is missing position information for several residues from 3 of its 4 domains. Namely, residues 1 to 26 from domain I, residues 346 to 367 from domain III, and residues 774 to 776 from domain IV. The *Molefacture* plugin from VMD was used to create peptide fragments for splicing with the LF crystal—a linear chain for residues 1 to 26, a helix for residues 346 to 367, and a linear chain for residues 774 to 776. The first 30 residues of the LF protein were then simulated using *NAMD* molecular dynamics and the *CHARMM[20]* forcefield at 300K and employing the *Generalized Born Implicit Solvent (GBIS[21])* with the remaining atoms held fixed, for ~10ns. Next, all atoms of the newly patched LF protein in complex with the MAPKK-2 peptide was simulated for ~0.5ns under the same conditions and subsequently minimized using *NAMD's* conjugate-gradient method. The result is shown in *figure 1*, without the peptide highlighted.

To generate the separate peptide and protein structures necessary for a blind docking search, the MAPKK-2 segment and patched LF protein were equilibrated separately using *NAMD* molecular dynamics. Running with constant NPT at 300K and 1atm with the CHARMM potential and TIP3P[22] explicit water, the LF protein was equilibrated for ~10ns. Similarly, the MAPKK-2 protein was simulated for ~5ns, and 100 sample protein conformations are chosen uniformly throughout the length of the trajectory. Due to large scale motion seen during the dynamics trajectory, the structures are aligned using all heavy atoms from residues 30 to 295 and 450 to 550. Analysis of the resulting structures continues in *Section 3*.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/31/2013 Terms of Use: http://spiedl.org/terms

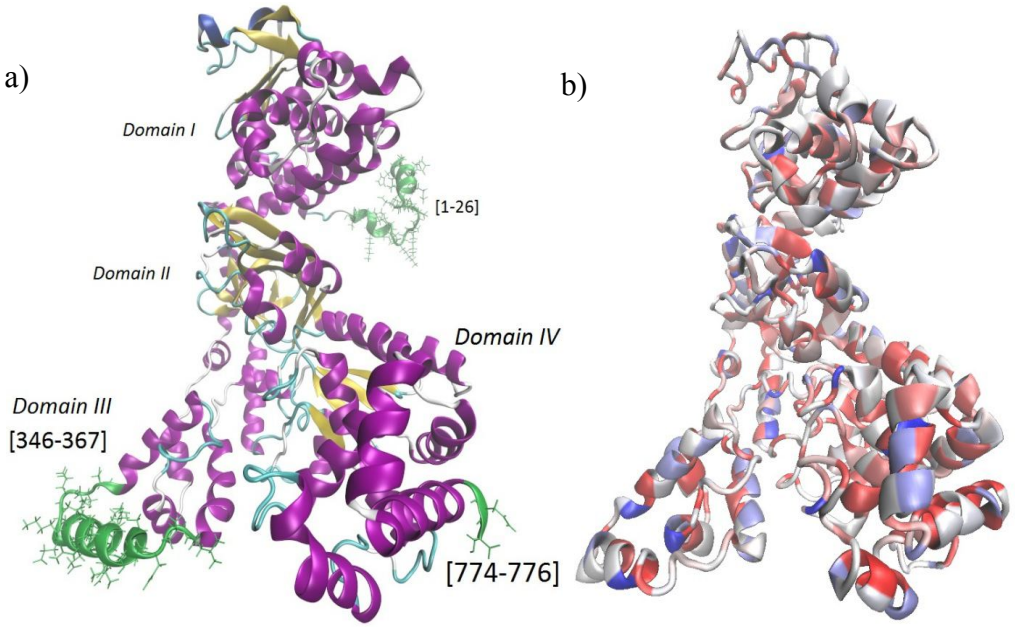

Figure 1. LF protein from the 1JKY structure. a) With missing residues labeled by residue number (GREEN) and b) colored using the Eisenberg hydrophobicity index (red=hydrophobic, blue=hydrophilic)

2.2 XPairIt Docking Protocol

The docking protocol used in this work generated ~50,000 configurations of the MAPKK-2 peptide in contact with the LF protein. The *NAMD Simulator[10]* and *PyRosetta[9]* interface to the *Rosetta Simulation Suite[23]* were used in conjunction with the XPairIt API as a means to improve sampling of peptide and protein flexibility during docking, and incorporate an improved energetic representation. To generate each docked configuration, a random peptide snapshot is drawn from the initial dynamics trajectory and placed in a simulation box with the equilibrated and minimized LF protein structure. Within the XPairIt API, atom positions are sent to *Rosetta* and the partners are randomly rotated around their centers of mass and moved into contact with one another until any pair of surface atoms of each partner is ~4.0Å apart. This effectively creates initial structures where the peptide is placed at thousands of different points on the protein's surface.

The *Rosetta DockingMCM* method and the *score12* score function are then used to perturb the peptide configuration, where small rotation and translation moves are sampled from a uniform distribution to determine the lowest energy structure. In addition to peptide moves, *Rosetta's DockingMCM* method samples sidechain rotamers of the peptide and protein. When the *DockingMCM* method is complete, the new atom positions are analyzed for structural and energetic properties in a bookkeeping step within XPairIt, and then sent to *NAMD*. Here, molecular dynamics is performed on the peptide and protein atoms within 15.0Å of the peptide, leaving other atoms fixed. This dynamics simulation is run at 300K using *GBIS*, including the solvent accessible surface area component (SASA), for 2.0ps, and then the structure is minimized for 3000 conjugate gradient steps. Finally, the structure's sidechains are again repacked with *Rosetta*, based on the *score12* scoring function, and this structure is exported as a PDB file. These steps are shown to improve sampling of the flexibility of the peptide and permit induced flexibility of the protein.[3, 8] To identify the best docked structure, the total energy and interface energy ($E_{INT} = E_{TOT} - E_{PRO} - E_{PEP}$) are used to rank the docking results.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/31/2013 Terms of Use: http://spiedl.org/terms

3. RESULTS

Here, we analyze and discuss the efficacy of the *XPairIt Docking Protocol* for docking peptides into a highly flexible protein, the Anthrax Lethal Factor. The structure of the MAPKK-2 peptide molecular dynamics trajectory is presented, and our best docked structure is investigated through a structural and energetic comparison to the 1JKY and 1PWV crystal structures.

3.1 MAPKK-2 Peptide Dynamic Structure Analysis

The sequence of the docked peptide, MLARRKPVLPALTINP, was chosen in previous studies to mimic the N-terminal end of mitogen-activated protein kinase kinase MAPKK2. This sequence has a very low helical propensity using the scale of Pace and Scholtz[24], largely due to the presence of three prolines. A map of secondary structure[25] during the course of the peptide trajectory shows very little meaningful structure, consisting primarily of a persistent 4 residue turn (T) around residues 10-13 (starting at Proline 10) and a transitory turn (T) approximately through residues 2-5. The rest of the structure is labeled as random coil. The radius of gyration is large, approximately 11.41 with a variance of 0.69. In short, the conformational picture of this peptide is that of a long, loose piece of string with a kink approximately 1/3 of the way from the end. This may be seen in Figure 2, where snapshots of the peptide backbone during the course of the simulation have been overlaid with the bound peptide backbone conformation from crystal structure 1JKY.

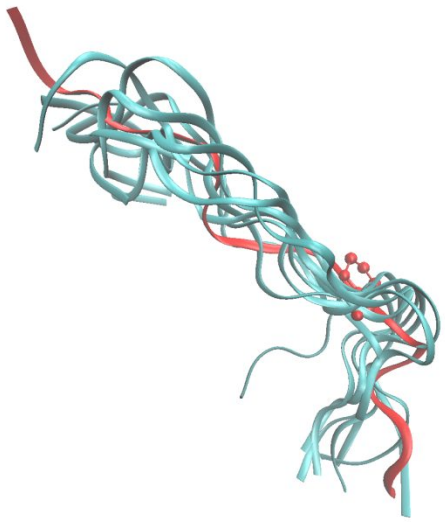

Figure 2. Overlay of peptide backbone structure during course of solvated dynamics trajectory (cyan). The backbone structure from the crystal structure 1JKY is shown in red, with Proline 10 shown in red CPK representation

3.2 MAPKK-2 Peptide and LF Protein Docking

Global docking (blind docking) of the MAPKK-2 peptide was performed with the equilibrated and complete structure of the LF protein, using the *XPairIt Docking Protocol*. Equilibration of docking partners and docking simulations proceeded as discussed in *Section 2*, and generated ~50,000 docked structures. Resulting configurations were first ranked by their total energy, then by their interface energy to determine the best docked structure. The best docked structure defined here is identified as rank #6 when ordered by interface energy and rank #1 when ordered by total energy.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/31/2013 Terms of Use: http://spiedl.org/terms

Table 1 lists the interface energies of the best docked structure over the course of the docking run, computed using the *score12* score function of *Rosetta*. After *Rosetta* docks the peptide and molecular dynamics is performed on atoms within 15 Å of the peptide, the interface energy increases. It is only after NAMD minimization and subsequent Rosetta minimization that the energy is actually improved, when compared to the beginning docked structure. Given that *Rosetta* repacks sidechain rotamers during the initial docking stage at the end of the protocol, it is likely that the NAMD molecular dynamics and minimization perturbs the system and sends it into a new minimum in configuration space.

Table 1. Interface energies of the best docked structure ranked by total energy, reported here in Rosetta *score12* format.

	Docked Score	NAMD Dynamics	NAMD Minimize	Rosetta Minimize	Rosetta Repack
Dock_446-134	-7.9875	-5.5521	-7.1058	-11.4303	-11.7799

We investigate the effect of *NAMD* molecular dynamics on the docked structure by first restating results from the equilibration of the LF protein which were reported in *Section 2*. Specifically, after a ~10ns molecular dynamics simulation in TIP3 explicit water, significant motion of domains III and IV was observed. This motion can be characterized as an opening of the binding pocket of the LF protein. A closer look at this area and analysis of domain III to domain IV distances reveals some effects of the dynamics during docking. Shown in *figure 3* is an overlay of the LF protein from the 1JKY crystal structure, after equilibration, and from the best docked structure. Measured distances between α-C's in ASN366 (domain III) and ASN732 (domain IV), as well as visual examination, show a large change during equilibration of the protein. There is also a significant decrease in both the ASN366-ASN732 distances and the overall structure of the LF protein after the docking simulation is completed, specifically in the loops of domain III. RMSD values relative to the 1JKY crystal and the equilibrated structure, and ASN distances are listed in *table 2*.
Also included in *table 2* are RMSD values of the LF protein in the best docked structure relative to the equilibrated structure, for residue Phe329, Glu676, and neighboring residues. According to Turk et al.[5], these residues exhibit ~3.0 to ~3.5Å shift upon binding of a 20-mer peptide to the LF protein. While values are lower in magnitude, there is similar motion captured during the docking simulation of MAPKK-2 peptide and the equilibrated LF protein.

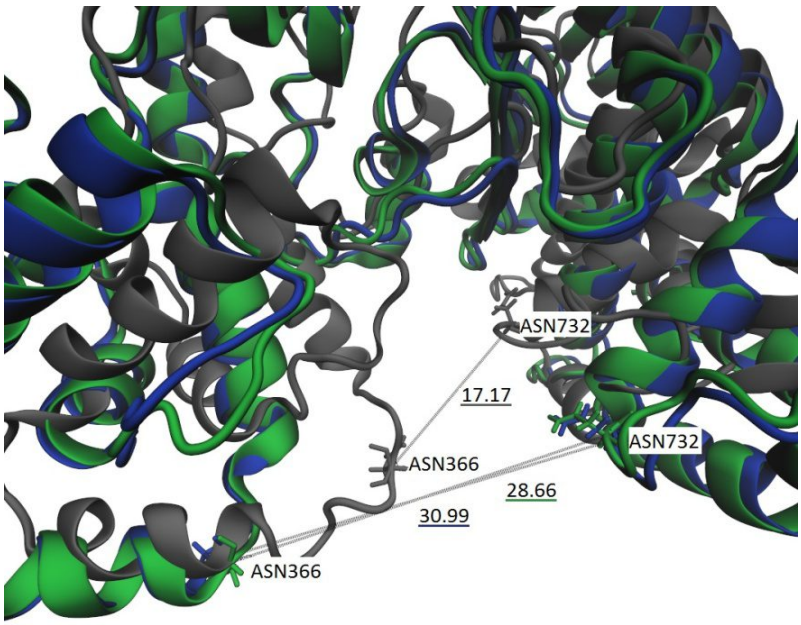

Figure 3. Overlay of the LF protein in the 1JKY crystal structure (GRAY), the equilibrated structure (BLUE), and the best docked structure (GREEN). Distances between ASN366 and ASN732 are labeled as 17.17, 30.99, 28.66 Å, respectively.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/31/2013 Terms of Use: http://spiedl.org/terms

Table2. RMSD values of the 1JKY crystal structure, equilibrated LF protein, and the best docked structure, from various comparisons. Structures are aligned by all heavy atoms from residues 30 to 295 and 450 to 550. Total RMSD is calculated by all heavy atoms from residues 30 to 730. All distances in Å.

	$RMSD_{CRYSTAL}$	$RMSD_{EQUIL}$	$RMSD_{EQUIL}$ [Asp328, Phe329, Leu330]	$RMSD_{EQUIL}$ [Val675, Glu676, Leu677]	ASN366 to ASN732
1JKY Crystal	0.0	--	0.0	--	17.17
Equilibrated LF Protein	5.613	0.0	--	0.0	30.99
Best Docked	5.546	0.959	0.953, 1.011, 1.845	0.483, 0.485, 0.584	28.66

Figure 4 provides an overlay of peptide-LF conformations from our simulation (where the peptide is the MAPKK-2 tail mimic and is shown in yellow), from the 1JKY crystal structure (where the peptide is the MAPKK-2 tail mimic and is shown in orange), and the 1PWV crystal structure (where the peptide is an optimized binder 20 residues long, typically denoted LF20, and is shown in red). We note that even in the absence of the catalytic Zn, which is expected to affect the electrostatic profile [6], the XPairIt protocol placed the peptide within the groove adjacent to the active site. It has been pointed out in the literature[5] that the 1JKY bound complex structure places the MAPKK-2 peptide too far from the active site. For this reason, we present the current results within the context of both the 1JKY structure and the 1PWV structure, making allowances for differences in the peptide sequence. Relative to both of the available crystal structures, residues 1 to 8 of the MAPKK-2 peptide in our docked configuration are folded forward and lay along domain III, rather than continuing up the groove towards domain II.

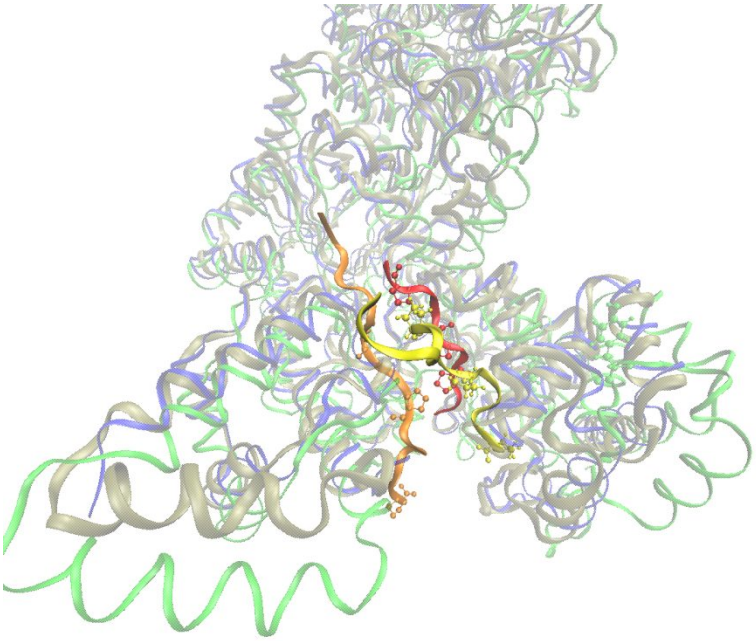

Figure 4. From global docking, the best docked structure using the *XPairIt Docking Protocol*, ranked by interface energy and total energy with Rosetta *score 12*. Best docked peptide conformation shown in *yellow*, 1JKY structure of Pannifer et al[4] peptide shown in *orange*, and the 1PWV structure LF20 peptide of Turk et al[5] shown in *red*. Prolines are added in CPK representation.

A key point in measuring the success of the docked structure is an assessment of distance between the active site and the S1 Proline adjacent to the scissile bond (where we follow the notation of Dalkas et al[7] to note residue placement along the length of the peptide). In the absence of Zn, we will measure the distance between the backbone carbon of MAPKK-2 peptide Proline 10 (or the LF20 peptide of structure 1PWV) and the epsilon nitrogen of the anthrax LF His686. His686 is chosen due to the fact that it coordinates the catalytic Zn[4]. In the 1PWV structure, which is nominally the 'correct'

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/31/2013 Terms of Use: http://spiedl.org/terms

configuration for cleavage, this distance is 4.52 Å. In the 1JKY structure, this distance is considerably larger at 13.82 Å. The XPairIt structure, while not as close as the 1PWV structure, is markedly better than 1JKY at 7.64 Å.

There has been much discussion within the literature regarding the importance of Tyr728 within the LF protein. While this residue is not in any position of interest in the 1JKY structure, in the 1PWV structure it does indeed play a palpable role, as the Tyr728 hydroxyl oxygen is 3.48 Å from the peptide Pro10 backbone carbon and a hydrogen bond is indeed a reasonable projection. This interaction is obviously missing in the XPairIt docked structure. A plot of the RMSD of this residue during the course of the dynamics of the LF protein alone shows an average drift of 6.33 Å away from its original location near the active site, and a total displacement by as much as 9.32 Å at some points during the run. To some extent this is due to the loss of a hydrogen bond connection between Tyr728 and Glu735. While this connection has been noted in the literature[6], in the original 1JKY crystal structure the oxygen-oxygen distance between the sidechains of these two residues is 3.48 Å, which is not indicative of a strong bond.

We turn to the question of what driving forces were felt by the MAPKK2 peptide in each of the two binding locations, the XPairIt conformation and the 1JKY conformation. Below, we present a breakdown by residue of contributions to the docking scoring function Score12, and the Coulomb interactions for the XPairIt conformation (Figure 5) and the 1JKY crystal structure (Figure 6). The central residues of the peptide (residue 9 and 10, plus various surrounding residues) provide primary interactions via the docking scoring function for both structures, particularly residue Leu 9. The scoring function total is quite similar for both complexes, with a total (in Rosetta units) of -4.67 for the XPairIt structure and -1.34 for the 1JKY structure. It is in comparing the Coulombic interaction that one begins to see the effect of the residues lining the groove, and their role in stabilizing the peptide. The 1JKY structure overall wins on Coulombic interaction, stabilizing the structure by -124.50 kcal/mol, versus -54.48 kcal/mol for the XPairIt structure. This effect is particularly evident in the first 6 residues of the peptide which line along the top of the groove in the 1JKY crystal structure (Figure 6). Despite its stabilizing influence, however, it is noteworthy that neither the LF20-LF complex (PDB:1PWV) nor the XPairIt structure demonstrate this interaction. Since both of these latter structures put the scissile bond in closer proximity to the active site than the 1JKY structure, it is possible that this interaction is useful but not necessary for cleavage.

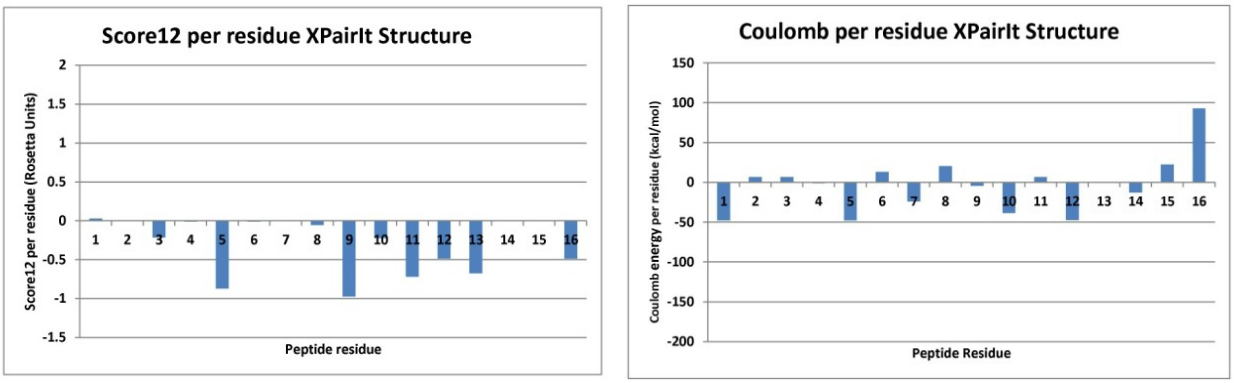

Figure 5. Docking score (LEFT) and Coulomb energy (RIGHT) per residue in final docked XPairIt structure

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/31/2013 Terms of Use: http://spiedl.org/terms

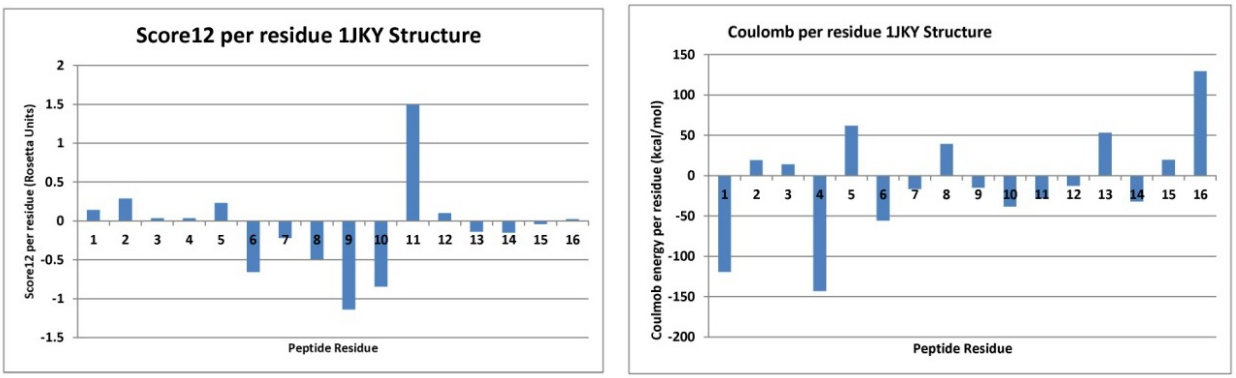

Figure 6. Docking score and Coulomb energy per residue in crystal structure 1JKY

The role of van der Waals interactions in stabilizing the interactions of MAPKK2-type peptides within the groove of LF has also been stressed within the literature[6] . Accordingly, we will choose to study the interactions of peptide residue LEU 9, which is seen in Figure 5 and 6 to provide some of the most favorable interactions along the length of the peptide, in addition to its proximity to the scissile bond. A breakdown of van der Waals score12 contributions for these two structures may be studied within the context of the reduced gradient density (RDG) tabulated by the NCIplot software[18, 19], and visualized below for the XPairIt structure (Figure 7) and the 1JKY structure (Figure 8).

The NCIplot software provides an electron density analysis similar in nature to the AIM or ELF analysis commonly used. With this tool, one is able to produce a plot of reduced electron density gradient (RDG) versus the electron density multiplied by the sign of the second Hessian eigenvalue. By focusing on regions of small RDG and small ρ, the researcher is able to quantify and characterize weak noncovalent interactions, separating hydrogen-bonds from van der waals interactions. Inclusion of the sign of the second Hessian eigenvalue allows for simple separation of attractions and repulsions. Due to the size of the protein-peptide complex, promolecular densities are used.

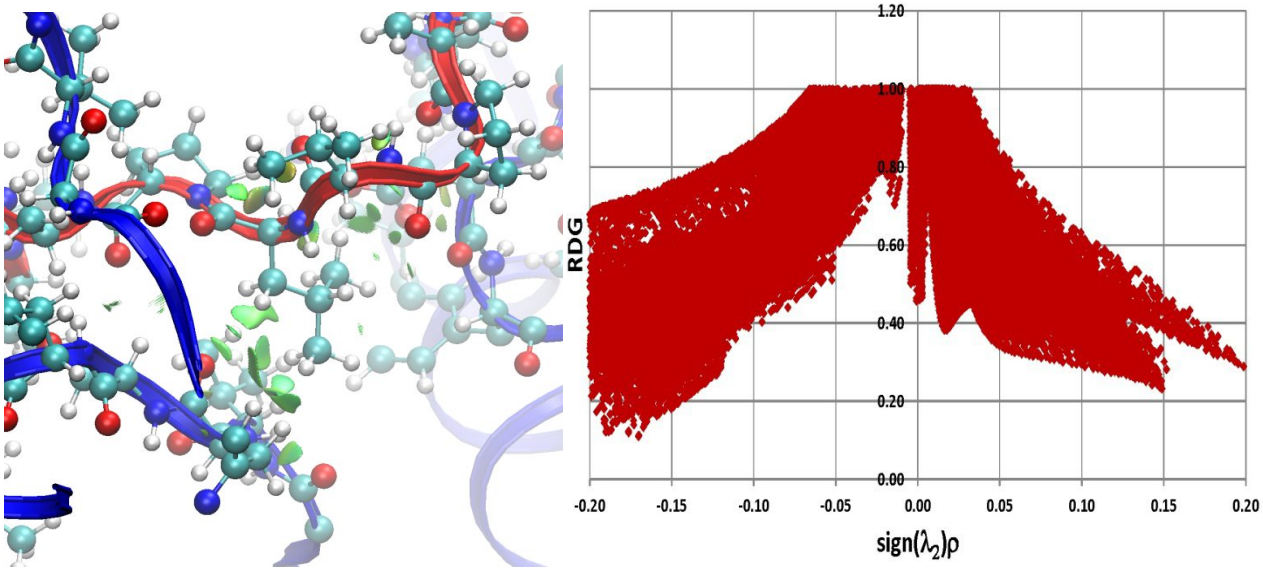

Figure 7. Isosurface and reduced gradient plot of environment surrounding Leucine 9 of XPairIt docked structure from NCIplot. In the isosurface representation of the RDG (LEFT), the LF protein backbone is depicted in blue ribbon representation, the Peptide backbone is depicted in red ribbon representation and the RDG isosurface is superimposed on the structure and is colored on a BGR scheme (Blue=attraction, Green=neutral, Red=repulsion). The RDG plot on the RIGHT shows minimal noncovalent interaction.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/31/2013 Terms of Use: http://spiedl.org/terms

An analysis of the isosurface plot superimposed on the XPairIt docked structure in combination with the RDG plot is very educational. No strong intermolecular interactions are visible (in agreement with the score12 analysis of the structure). The areas of greatest overlap (largest area of visible isosurface) correspond on a per residue basis almost exactly to the per residue score12 contribution plotted in Figure 5. Figure 7 is a closeup of the MAPKK2 peptide Leucine residue 9 interactions with the LF protein. From the RDG plot on the right we see no strong interactions, only a weak spike about $\rho=0$ in the center of the plot. This corresponds to the interactions pictured in the isosurface image on the left, which are weak van der Waals.

Results for the 1JKY structure, in comparison, also consists primarily of weak interactions. Some weak steric clashes are evident in the light areas of yellow which appear on the RDG isosurface. Again, no hydrogen bonds are visible. Peptide Leucine 9 is again a major contributor to the score12 per residue breakdown, as can be seen in Figure 8. Here we see more visible surface in the isosurface plot than in the XPairIt structure, indicating more area of interaction, and a strong spike about $\rho=0$ in the RDG plot on the left. This again corresponds to van der Waals, but indicate a stronger interaction than in the XPairIt structure.

These trends are in qualitative agreement with the score12 function, which has assigned a value for -2.528 Rosetta units for the van der Waals attraction component of the interactions of Leu9 in the XPairIt structure, and a value of -3.456 Rosetta units for the 1JKY structure.

As a counter balance to these weak interactions, Figure 9 depicts similar information for the LF20-LF complex from 1PWV. Here we see additional peaks shifted away from the $\text{sign}(\lambda_2)\rho=0$ axis of the RDG plot, representing stronger attractions (to the left of the axis) and repulsions (to the right of the axis). One of these attractive interactions in particular is highlighted in the isosurface plot on the left, a hydrogen bond involving the OH group of Tyr728.

This type of iterative analysis, analyzing interactions residue by residue within XPairIt and characterizing the nature and strength of these interactions represents a valuable tool in understanding the strengths and weaknesses of a proposed complex structure, and may guide the investigator toward finding missing interactions.

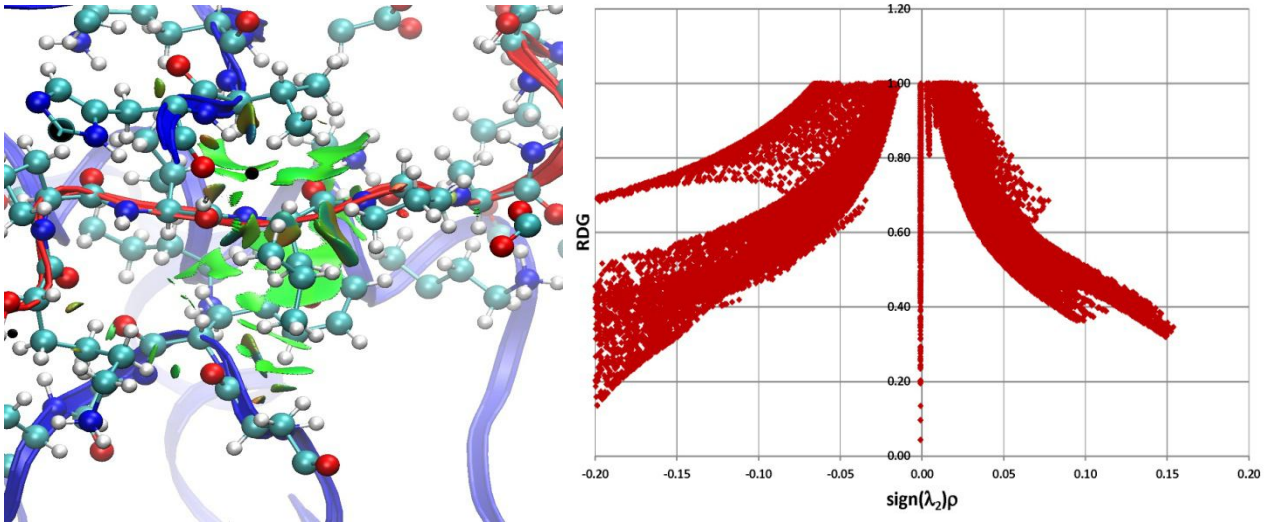

Figure 8. Isosurface and reduced gradient plot of environment surrounding Leucine 9 of the 1JKY structure from NCIplot. In the isosurface representation (LEFT), the LF protein backbone is depicted in blue ribbon representation, the Peptide backbone is depicted in red ribbon representation, and the RDG isosurface is superimposed on the structure and is colored on a BGR scheme (Blue=attraction, Green=neutral, Red=repulsion) The RDG plot on the RIGHT shows some weak van der Waals interaction.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/31/2013 Terms of Use: http://spiedl.org/terms

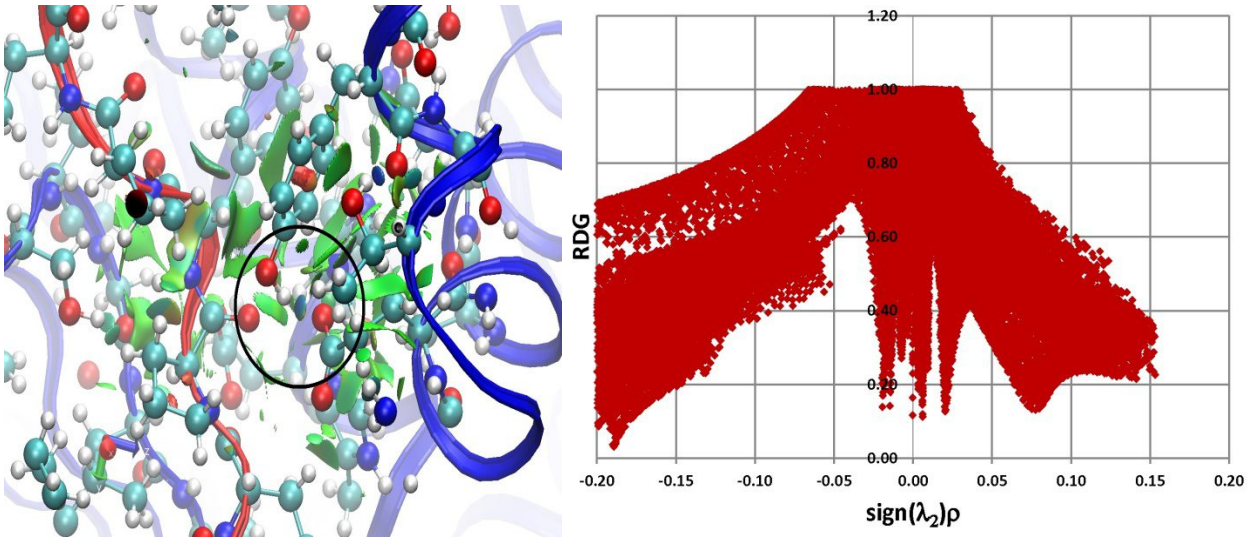

Figure 9. Isosurface and reduced gradient plot of environment surrounding LF Tyr728 of the 1PWV structure from NCIplot. In the isosurface representation (LEFT), the LF protein backbone is depicted in blue ribbon representation, the Peptide backbone is depicted in red ribbon representation, and the isosurface is superimposed on the structure and is colored on a BGR scheme (Blue=attraction, Green=neutral, Red=repulsion). The RDG plot on the RIGHT shows an increase in both attractive and repulsive noncovalent interactions visible as sharp spikes to the left and right of the $\text{sign}(\lambda_2)\rho=0$ axis.

4. CONCLUSION

We have presented an analysis of results of a multi-method, blind docking simulation of the complex formed between peptides and the anthrax lethal factor (LF), and compared these results to experimental crystal structures of two known complexes. Analysis of the noncovalent interactions involved in the complexes of the MAPKK2-mimic peptides, both our own proposed structure and the 1JKY crystal structure of Pannifer et al, showed very weak van der Waals interactions dominating. The 1JKY structure shows additional electrostatic stabilization for the MAPKK2 peptide within the upper portion of the LF groove. However, this interaction is missing in both the XPairIt docked structure of the MAPKK2 peptide, and also in the LF20-LF complex structure of Turk et al. Both the XPairIt structure and the LF20-LF structure placed the scissile bond of the peptide markedly closer to the cleavage site. We demonstrated the usefulness of the reduced density gradient (RDG) for characterizing the interactions within each of these complexes and proposed it may prove useful in validation of docked structures and as an aid for finding missing interactions.

5. ACKNOWLEDGEMENTS

This project receives support from the *Defense Threat Reduction Agency Joint Science and Technology Office for Chemical and Biological Defense* (Grant no. BRCALL08-Per3-P-2-0028). *NAMD* was developed by the Theoretical and Computational Biophysics Group in the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign. Images were generated using *VMD*. We gratefully acknowledge helpful discussions with Dimitra Stratis-Cullum, Deborah Sarkes, Amethist Finch, and Bryn Adams of the US Army Research Laboratory, as well as Alex Balboa of the US Army ECBC and Weitao Yang of Duke University.

6. REFERENCES

[1] N. London, D. Movshovitz-Attias, and O. Schueler-Furman, "The Structural Basis of Peptide-Protein Binding Strategies," Structure, 18, 188-199 (2010).

[2] D. N. Stratis-Cullum, J. M. Kogot, M. S. Sellers *et al.*, "Development of bacterial display peptides for use in biosensing applications," Proceedings of the SPIE. 8358, 83580A (2012).

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/31/2013 Terms of Use: http://spiedl.org/terms

[3] M. Sellers, and M. M. Hurley, "In silico design of smart binders to anthrax PA," Proceedings of the SPIE. 8358, 835807-835807-9 (2012).
[4] A. D. Pannifer, T. Y. Wong, R. Schwarzenbacher *et al.*, "Crystal Structure of the Anthrax Lethal Factor," Nature, 414(8 November, 2001), 229-233 (2001).
[5] B. E. Turk, T. Y. Wong, R. Schwarzenbacher *et al.*, "The structural basis for substrate and inhibitor selectivity of the anthrax lethal factor," Nature Structural and Molecular Biology, 11(1), 60-66 (2004).
[6] R. Hong, A. Magistrato, and P. Carloni, "Anthrax Lethal Factor Investigated by Molecular Simulations," Journal of Chemical Theory and Computation, 4, 1745-1756 (2008).
[7] G. A. Dalkas, A. Papakyriakou, A. Vlamis-Gardikas *et al.*, "Insights into the anthrax lethal factor-substrate interaction and selectivity using docking and molecular dynamics simulations," Protein Science, 18, 1774-1785 (2009).
[8] M. S. Sellers, and M. M. Hurley, "Application of the XPairIt API for Peptide Docking and Analysis," submitted to Journal of Computer Aided Molecular Design.
[9] S. Chaudhury, S. Lyskov, and J. J. Gray, "PyRosetta: a script-based interface for implementing molecular modeling algorithms using Rosetta," Bioinformatics, 26, 689-691 (2010).
[10] J. C. Phillips, R. Braun, W. Wang *et al.*, "Scalable molecular dynamics with NAMD," Journal of computational chemistry, 26, 1781-1802 (2005).
[11] I. Antes, "DynaDock: A new molecular dynamics-based algorithm for protein-peptide docking including receptor flexibility," Proteins: Structure, Function, and Bioinformatics, 78, 1084-1104 (2009).
[12] J.-H. Lin, A. L. Perryman, J. R. Schames *et al.*, "Computational drug design accommodating receptor flexibility: the relaxed complex scheme," Journal of the American Chemical Society, 124, 5632-5633 (2002).
[13] A. H. Elcock, D. Sept, and J. A. McCammon, "Computer Simulation of Protein-Protein Interactions," The Journal of Physical Chemistry B, 105, 1504-1518 (2001).
[14] M. Nowosielski, M. Hoffmann, A. Kuron *et al.*, "The MM2QM tool for combining docking, molecular dynamics, molecular mechanics, and quantum mechanics," Journal of computational chemistry.
[15] N. Okimoto, N. Futatsugi, H. Fuji *et al.*, "High-performance drug discovery: computational screening by combining docking and molecular dynamics simulations," PLoS Computational Biology, 5, e1000528 (2009).
[16] D. Seeliger, and B. L. De Groot, "Ligand docking and binding site analysis with PyMOL and Autodock/Vina," Journal of computer-aided molecular design, 24, 417-422.
[17] M. A. Lill, and M. L. Danielson, "Computer-aided drug design platform using PyMOL," Journal of computer-aided molecular design, 25, 13-19.
[18] E. R. Johnson, S. Keinan, P. Mori-Sanchez *et al.*, "Revealing Noncolvalent Interactions," Journal of the American Chemical Society, 132, 6498-6506 (2010).
[19] J. Contreras-Garcia, E. R. Johnson, S. Keinan *et al.*, "NCIplot: A Program for Plotting Noncovalent Interaction Regions," Journal of Chemical Theory and Computation, 7, 625-632 (2011).
[20] A. D. MacKerell Jr, D. Bashford, M. Bellott *et al.*, "All-atom empirical potential for molecular modeling and dynamics studies of proteins," The Journal of Physical Chemistry B, 102, 3586-3616 (1998).
[21] W. C. Still, A. Tempczyk, R. C. Hawley *et al.*, "Semianalytical treatment of solvation for molecular mechanics and dynamics," Journal of the American Chemical Society, 112, 6127-6129 (1990).
[22] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura *et al.*, "Comparison of simple potential functions for simulating liquid water " Journal of Chemical Physics, 79, 926-936 (1983).
[23] A. Leaver-Fay, M. Tyka, S. M. Lewis *et al.*, "ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules," Methods in Enzymology, 487, 545-574.
[24] C. N. Pace, and J. M. Scholtz, "A Helix Propensity Scale BAsed on Experimental Studies of Peptides and Proteins," Biophysical Journal, 75, 422-427 (1998).
[25] D. Frishman, and P. Argos, "Knowledge-based protein secondary structure assignment," Proteins: Structure, Function, and Bioinformatics, 23, 566-579 (2004).

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/31/2013 Terms of Use: http://spiedl.org/terms

NO. OF COPIES	ORGANIZATION
1 (PDF)	DEFENSE TECHNICAL INFORMATION CTR DTIC OCA
1 (PDF)	DIRECTOR US ARMY RESEARCH LAB IMAL HRA
1 (PDF)	DIRECTOR US ARMY RESEARCH LAB RDRL CIO LL
1 (PDF)	GOVT PRINTG OFC A MALHOTRA

ABERDEEN PROVING GROUND

1 (PDF)	DIR USARL RDRL WML B M HURLEY